

METHOD FOR OPTIMIZING THE POSITIONS OF STANDARDS ON MALDI TARGETS FOR MASS SPECTROMETRY

FIELD OF THE INVENTION

The invention relates to the field of molecular mass measurement by time-of-flight mass spectrometers where the ionization of analyte substances is produced by matrix-assisted laser desorption (MALDI). Methods for the correction of flight time values based on the use of external standards are provided, and in particular methods for determining, given a MALDI target, the optimal positions of mass calibration standards are described.

BACKGROUND

Among the methods for ionization of macromolecular substances from sample supports, matrix-assisted desorption by a laser flash (MALDI) has found the widest acceptance. In preparation for ionization by MALDI, sample molecules are spotted on a support (target) in or on a layer of low-molecular weight matrix substance. Different samples are generally present together as separate spots on the same MALDI target. In the mass spectrometer, a laser pulse of a few nanoseconds duration, focused onto the target surface, vaporizes a small amount of the matrix in a quasi-explosive process, thus transferring the sample molecules into an initially small vapor cloud. The vapor cloud expands into the vacuum of the mass spectrometer, accelerating the matrix molecules and ions through its adiabatic expansion, as well as the sample molecules and ions through viscous entrainment. The ions are then accelerated with electrical fields of around 10 to 30 kilo-electronvolt (keV) towards the flight path of the mass spectrometer and subsequently detected with high time resolution at the end of the flight path. The mass-to-charge ratios are determined from the flight times. This type of ionization typically supplies only singly charged ions, thus the term "mass determination" will be used herein instead of the more precise term "determination of the mass-to-charge ratio".

As mentioned above, MALDI-TOF mass spectrometric analysis (Karas, Hillenkamp Anal. Chem. 60 (1988) 2299) requires that samples are deposited on a plate or target (see Figure

1). State-of-the-art MALDI mass spectrometers are designed and tuned to provide high mass accuracy and high spectral resolution. Nevertheless, optimal mass precision requires a calibration step in addition to the default calibration of the instrument. For instance, a spectrometer may show a mass determination offset by 0.3 Da with respect to theoretical values. If no further
5 external or internal calibration is performed, this systematic error cannot be compensated for. By contrast, if at least one sample on the plate contains a mixture of compounds the masses of which are known, this 0.3 Da deviation is easily detected and corrected.

Two techniques are commonly used to calibrate MALDI-TOF mass spectra: (1) internal calibration, where at least one known compound (standard) is added to every sample such
0 that each spectrum contains the known mass(es) and (2) external calibration, where standards and samples are separated on the MALDI target (e.g., in separate spots). An exemplary internal calibration method is described by Köster, et al. in U.S. Patent 5,886,345. There are cases however, where addition of reference standards to a test sample may not be desired. For example, the standards may fall outside the ideal mass range, be incompatible with the test sample, or be
5 present in a comparatively large amount relative to the test analyte in the sample. External calibration is standard practice in the art and commercially available MALDI targets may have a defined set of standard positions on one or more side of the target. While such external reference masses and calibration methods have found widespread use in MALDI-TOF methods, various factors, e.g. arising from spatial distance between sample and standard positions on the target, or
10 from time lag between the measurement of the standard and a given sample spot, still affect the accuracy of observed masses. Such factors result in an offset or drift that occurs in measured readings over time and/ or distance on the target surface. In view of the many sources of potential error in mass determination and the precision required, there is a particular need for a calibration method that does not rely on pre-determined assumptions about the nature or source of
25 the experimental error. It is an object of the invention to improve mass calibration methods by optimally positioning standards on a target to minimize factors that cause drift in measured mass readings.

SUMMARY OF THE INVENTION

30 It is an object of the present invention to provide a method to optimally position standards on a target in view of external calibration of MALDI-TOF mass spectra. For a given mass spectrometer and MALDI target set, a number N of positions on the target are usually chosen to be dedicated to external standards. As such these positions are not available for sample

deposition. The greater the number N , the better the calibration, to the expense, however, of a decreased throughput in the number of samples analyzed. There is thus a balance to be found between the number of spots dedicated to standards and the remaining number of spots dedicated to samples. For a fixed number N , however, the positioning of the standards on the target, and the order in which the positions are analyzed, influence the results of the calibration. The present invention thus provides a method which optimizes these two parameters and which therefore maximizes the accuracy of the external calibration.

In the present invention, a calibration method is used to test, potentially all, the positions on a MALDI target as suitable for standard deposition, in order to optimize the positioning of the standards which will subsequently be used in the analysis of samples. This method improves accuracy over arbitrary standard position selection.

In one aspect of the present calibration method, samples containing a mixture of compounds of known masses (standards) are spotted onto a target (hereinafter, "calibration target") and analyzed by mass spectrometry. Preferably, every position on the calibration target is loaded with the mixture to allow testing at each position. For purposes of explanation, we assume that the MALDI target contains a total of P positions, of which N will be used for standard deposition in the final layout once the optimal positions have been identified (leaving $(P-N)$ positions available for samples).

According to the method of the invention, the data generated by the mass spectrometric analysis of the calibration target is processed in the following way: for every combination of N positions on the calibration target:

(i) the $(P-N)$ positions that do not belong to the current N positions are calibrated by applying a calibration procedure based on the data acquired from the current N positions; and

(ii) the performance of the current N positions for calibrating the $(P-N)$ positions is computed according to a given criteria.

Finally, the method involves selecting the combination of N positions that gives the best performance, for calibrating the $(P-N)$ positions, to be used in the final target layout.

In one aspect, the invention provides a method to optimally position standards on a target, comprising the steps of:

- (1) providing a target having P positions containing a standard and generating mass spectrometric data from these P positions;
- (2) providing a combination of N positions on the target ($N < P$), and calibrating the mass spectrometric data acquired from the $(P-N)$ positions that do not belong to the current N

positions by applying a procedure based on the mass spectrometric data acquired from the current N positions;

(3) computing the performance of the current N positions in calibrating the (P-N) positions according to a given criterion;

(4) carrying out steps (2) and (3) for a plurality of combinations of N positions on the target; and

(5) selecting a combination of N positions that gives a desired performance as a combination that may be used for positioning the standards in the final target layout.

When referring to a combination of N positions of mass standards, the term "performance" refers to the ability of the combination of N positions to calibrate the positions used in a given analysis procedure, such as the P-N positions. In a preferred embodiment, the criterion applied for defining the performance of a combination of N positions is the standard deviation of the statistical distribution of mass differences between the theoretical, exact masses of the standards at the (P-N) positions and the experimentally measured masses at these positions, after calibration by reference to the N positions. While the positions used for determining calibration performance preferably equal P-N, it will be appreciated that performance can be determined for any desired position usage. As such, the total positions used when referring to performance need not necessarily be the total available positions for sample deposition on a target.

The desired performance will generally be the best performance, such that the combination of N positions selected gives optimum performance in calibrating a MALDI target. For example, the 1, 2, 5 or 10 best performing combinations of N positions can be selected. However, the desired performance may also be based on any other suitable criteria, such as better than the average performance computed for the plurality of N combinations, or performance within the top 30%, 20%, 10%, 5% or 1% of N combinations tested, etc.

The number N of positions dedicated to external standards in the final target layout can be any suitable number. N may be 1, or generally more, preferably at least 2, 3, 4, 5, 6, 7, 10, 15, 25, 50 or 100 positions. Most preferably, N is between 1 and 25.

Preferably the method, or more particularly steps (2) and (3), are carried out for every combination of possible N positions on a target.

The target is most preferably a MALDI target. MALDI targets are well known in the art and can be obtained commercially.

The invention provides a method that, given a number N of positions dedicated to standards, allows for determining the optimal positions of the standards. In this method, every position of one MALDI target ("calibration target") contains standards and preferably no unknown samples. If N is the number of positions that will contain a standard in the final layout, the best N -tuple of positions is determined by testing every N -tuple for a given criterion. The performance of the N -tuple used by the given calibration procedure is evaluated on the drift in observed readings for the other ($P-N$) positions (also containing known standards in the calibration target).

In further preferred embodiments, more than one calibration target are used to increase confidence in the optimal solution of combinations of N positions, especially where a common combination of N positions is desirable for a number of different targets. Furthermore, more than one calibration target may be used on different MALDI instruments, such that an optimal position for standards may be provided which can be used on different machines or types of machines.

Alternatively, different MALDI instruments may be calibrated independently, such that the final target layout may be different for each instrument.

In yet another embodiment, each MALDI target is calibrated individually. MALDI targets are commonly re-used in the art, and an optimal standard layout may thus be determined for each different target. This will allow physical differences between targets, such as non-perfect planarity, to be taken into account in the calibration.

In another embodiment, individual MALDI target are calibrated individually for a specific instrument, such that, when using a different instrument with the same target, or when using the same instrument with a different target, another set of standard positions is used.

In further embodiments, not every position (e.g., less than P) are used on at least one of the calibration targets. It will also be appreciated that in preferred embodiments, performance is not computed for every N -tuple of positions. That is, a subset every N -tuple of positions is tested.

In further embodiments, a heuristic method is used to generate possible combinations of N positions. In preferred aspects, a heuristic method is selected from the group consisting of simulated annealing, genetic algorithm, taboo search, hill climbing, conjugate gradient and ant systems.

In preferred embodiments, the combinations of N positions to be tested are reduced by using geometrical symmetry. In a typical embodiment, the MALDI target surface is divided into a number of sectors. For example, the MALDI target surface may be divided into N sectors of sensibly equal areas, each containing one of the N positions for standards, e.g. in the center of the sector.

In another embodiment, a subset M of the N positions is fixed beforehand and the testing for optimizing standard positions is conducted only on the remaining (N-M) positions.

In preferred embodiments, the criterion for performance of the calibration resulting from an N-tuple of positions is the standard deviation of the statistical distribution of the mass

5 differences between the theoretical, exact masses of the standards at the (P-N) positions and the experimentally measured masses at these positions, after calibration by reference to the N positions. Alternatively, in any of the methods of the invention, the criterion for performance of the calibration resulting from a N-tuple of positions may be the skewedness of the mass differences distribution.

10 The invention also provides methods of using the information obtained while optimally positioning standards. Targets having optimal standard positions can be used to calibrate the mass of a test molecule, preferably from a position not containing standards. The invention thus further comprises use of a MALDI target comprising a combination of N standard-containing positions, which combination of positions is obtained according to the methods of the invention.

15 The methods further comprise obtaining a mathematical transformation which can be used to calibrate a mass in a nonstandard- containing position. The invention also encompasses calibrating a mass using a MALDI target arranged according to the invention.

The invention also provides novel calibration methods. Said methods may be used in methods to optimally position standards on a target as well as in methods of using targets
20 comprising optimally positioned standards and the mathematical transformations obtained therefrom.

The methods of the invention may provide a single mathematical transformation for a combination of N positions for use in calibrating other positions, or may provide a plurality of mathematical transformations for a combination of N positions. Also encompassed is an
25 embodiment where the calibration procedure improves measurements made in positions not containing standards by using a mathematical transformation obtained from the closest standard-containing position.

In preferred aspects, the calibration procedure makes use of polynomial transformations computed by regression from the mass spectra obtained from standard-containing positions.
30 Preferably, P-N positions are calibrated using a polynomial transformation obtained from N positions.

Preferably, the mass spectrum of the standard-containing position is acquired prior to those of the positions to be calibrated.

In further preferred embodiments, the invention provides methods that reduce the drift that occurs in time during mass acquisition. In this aspect, the invention provides methods of calibrating a sample positioned on a target, wherein the spectra of the samples located at positions on the target farther away from the standard-containing position they depend on for calibration are acquired before those of the positions closer to the standards.

In further preferred methods, the spectrum of standards is re-acquired periodically to reduce temporal errors during calibration.

The calibration procedure can be carried out according to known methods. Preferred methods are provided herein. Further described herein are methods of using, for sample analysis, the optimal layout obtained by the method of the invention. In one aspect, the invention encompasses a method of detecting the mass of an analyte of interest and positioning standards on a MALDI target, comprising defining a plurality of spatial zones on a MALDI target, each zone comprising a standard, wherein each standard is positioned substantially in the middle of, and preferably equidistant from, the outer bounds of its respective zone.

The invention thus provides a correction parameter, or mathematical (polynomial) transformation, to calculate the exact masses of analyte ions, which uses the measured mass of the standard ions. Thus, the masses of ions from added known substances are measured, and the masses of the analyte ions are corrected by extrapolation.

DESCRIPTION OF THE FIGURES

Figure 1 shows an schematized exemplary MALDI target.

Figure 2 shows the scheme of a calibration procedure used to calibrate analytes based on external references (standards). Figure 2A shows an example of a target layout with the standard positions in bold. The other positions are labeled with the number of the standard used for their calibration. Figure 2B shows an illustration of the transformation function: 3 experimental masses are measured and the affine transformation $y=ax+b$ is used to calibrate. Higher degree polynomials are possible depending on the number of known masses.

Figure 3 shows a series of graphs demonstrating error frequencies after mass calibration of four MALDI targets (horizontal unit is ppm). Setting (C) is superior to settings (A) and (B) as error frequencies are, for all 4 MALDI targets, more centered around 0. See Detailed Description for the exact definition of settings (A), (B) and (C).

Figure 4 shows an optimal layout for a MALDI target which uses 5 positions for external standards. Each standard is used in the calibration of a region of identical area. Within each area,

the acquisition is conducted, after measuring the mass of the standard in the center, by starting from the sample position most remote from the standard position and by subsequently following a spiral progression towards the standard containing center position (illustrated in the bottom left area).

5 Figure 5 presents the details of a calculation performed in order to calibrate the masses obtained from one position, based on standards deposited onto another MALDI plate position.

DETAILED DESCRIPTION

As used herein, a "sample" refers to any mixture of chemical compounds, preferably proteins or peptides of interest to be analyzed by mass spectrometry.

10 As used herein, a "standard" is understood to mean any known chemical compound, preferably a protein or peptide. A "standard" may also be a mixture of known chemical compounds or proteins or peptides. The masses measured from the standards are known. A "standard" or "standards" often comprises a plurality of peptides or proteins of known mass, preferably a standard comprises at least 1, 2, 3, 4, 6, 8 or 10 peptides or proteins of known mass.

15 As used herein, a "target" or "plate" is a substrate or surface on which samples are deposited after appropriate preparation and prior to MALDI mass spectrometry analysis. The methods described herein are compatible with any known preparation procedure for sample deposition on the MALDI target (e.g. Kussmann, et al., J. Mass Spectrom. 32 (1997) 483-493). As targets, metal plates with or without anchor technology (as described in US patent 6,287,872)
20 can be used. A schematic MALDI target is shown in Figure 1.

As used herein, a "well" or "spot" refers to a position on a target. Spots typically define a position where a compound or sample is or can be deposited. An example of a plurality of spots is shown in Figure 1. In preferred embodiments, each spot is defined by a hydrophilic anchor (U.S. Patent 6,287,872). This hydrophilic anchor is placed on the hydrophobic background of the
25 sample support such that the sample (generally in aqueous solution) is precisely attracted to that particular position.

As used herein, a "calibration" is referred to as a transformation aimed at improving the precision of a measurement, preferably a measurement of mass. As used herein, a "calibration procedure" refers to any procedure that, given a set of standard-containing positions, defines a
30 transformation aimed at improving precision of the mass determination for an analyte of interest that is preferably not at a standard-containing position (see Figure 2). As used herein, a "calibration target", is a MALDI target prepared for the purpose of the method of the invention,

wherein standards are deposited both at the N positions for standards and at the (P-N) positions for samples.

Two techniques are commonly used to calibrate MALDI-TOF mass spectra: (1) internal calibration, where at least one known compound (standard) is added to every sample such that each spectrum contains the known mass(es) and (2) external calibration, where standards and samples are separated (e.g., in separate spots). The spectra acquired from samples are calibrated by using the spectra acquired from the standards.

Disclosed herein is a new method of optimizing the position of standards in order to perform calibration. The method is most advantageously used in external calibration methods. In particular, given a number N of positions dedicated to standards in a calibration procedure, the invention provides a method to find the standard positions that ensure the best calibration of the mass spectra acquired from the samples.

As used herein, MALDI spectrometers include well-known instruments such as the MALDI-TOF shown in James, P. ed., 2000: *Proteome Research: Mass Spectrometry*, Springer, Berlin. MALDI-TOF spectrometers measure mass data from samples deposited on a target (see Figure 1). Drift or offset in measured masses may result from several factors, however, spatial and temporal factors are the most influential. Spatial factors result from the sample position on the target relative to the standard position. Such factors may be explained by a non-uniform electric field in the neighborhood of the target or by the fact that the target is not perfectly planar. Temporal factors result from the instrument drift occurring during the time delay between two acquisitions and can be observed for two measurements made at the same position on the target but at two different points in time.

We assume a calibration procedure to rely on a set of N positions containing standards of known masses. The invention may be described by the following calibration procedure (see Figure 2). Once the N standard positions have been fixed on the target, a mathematical transformation is computed for each of them to map the measured masses at these positions to the corresponding known masses contained in the standard deposited. For instance, for $m \geq 2$ masses in the standard, an affine application to transform the experimental masses may be found by linear regression which minimizes the sum of squared differences between experimental and theoretical standard masses. Subsequently, for each measured sample, based on its position on the target, the closest standard position is determined, and the associated affine transformation is used in calibrating the measured sample masses. See Figure 5 for an example of computation.

To illustrate the impact of the standards positions, we present an example with a standard mixture of four peptides. The mixture is applied to each position on four 384-positions containing

targets. Three settings for calibration are compared in Figure 3. In setting (A), only one position is used for the calibration. The position containing the standard is measured first and, according to an affine transformation obtained by linear regression, the measured masses at the other positions are calibrated. In setting (B), five standard positions are used: one at the center of the target and one at each corner. In setting (C), five standard positions are used as well but, in this case, the optimal standard positions discovered by the method of the invention are used. In settings (B) and (C), the calibration procedure described above has been applied (affine calibration, see figure 5 for an example). The optimal standard positions as used in setting (C) involve the segmentation of the target surface into five virtual sections of substantially equal areas, each of them containing a position for standard in its center (see Figure 4).

Thus, in one aspect, the present invention provides a method to optimally position standards on a target in view of external calibration of MALDI-TOF mass spectra.

First, a mixture of compounds of known masses is deposited at specific positions on a MALDI calibration target. In preferred embodiments, every position on the calibration target is used, thereby permitting an exhaustive analysis of every possible standard position. Thus, in one embodiment, 384 positions are used. However, it will be appreciated that the mixture may be applied to any other suitable number of positions, depending on the number of positions for which optimal standard placement is being investigated.

Thus, in one example, a MALDI target can be said to contain a total of P positions. Standards will be applied to N positions in the desired final target layout. N can be any suitable number; in general at least 1, preferably at least 4, or at least 5, 6, 8, 10, 20, 50 or 100 depending on the total number of positions P and on the quality of calibration that is desired.

According to the invention, after mass spectrometry analysis of the calibration target, multiple combinations of N positions are then tested *in silico* in order to determine their relative performances and thereby determine the optimal combination of positions for the standards. Preferably, every combination of N positions is tested.

The (P-N) positions that do not belong to the current combination of N positions (i.e., the combination being tested) are calibrated based on the mass spectrometric data from the current N positions. For this purpose, any calibration procedure may be used, as described herein.

The performance of a current combination of N positions is then computed by applying a given criterion. The criterion applied for defining the performance of a combination of N standard positions is for example the standard deviation of the statistical distribution of mass differences between the theoretical, exact masses of the standards at the (P-N) positions and the

experimentally measured masses at these positions, after calibration by reference to the N positions. The (P-N) positions also contain the mixture of standards on the calibration plate. The combination of N positions that gives the best performance, e.g., the lowest standard deviation, is then selected for positioning the standards.

5 Preferably, the method above is applied to several targets to obtain a better confidence in the resulting optimal set of standard positions. Alternatively, it is also possible to apply the method where the standard mixture is not applied to every position of the calibration target.

In a preferred embodiment, a heuristic algorithm is applied to generate possible combinations of N positions. Thus, in this embodiment, every combination of N positions may
0 not be tested. Such algorithms include: simulated annealing, genetic algorithm, taboo search, hill climbing (Michalewicz, Z. and Fogel, D.B. (2000) *How to Solve It: Modern Heuristics*, Springer, Berlin), and ant systems (Dorigo, M., Maniezzo, V. and Colorini, A. (1996) *The ant system: optimization by a colony of cooperating agents*, IEEE Transactions on Systems, Man, and Cybernetics-Part B, 26:1-13).

15 The calibration procedure used in the methods of the invention can be any of the known calibration procedures in the art. In a preferred aspect, the calibration procedure comprises the following steps:

- 20 1. Computing a polynomial transformation by regression (Spiegel, M.R. (1972) *Theory and Problems of Statistics*, McGraw-Hill, New-York) for each position in the combination of N positions. This requires a number of compounds in the standard mixture that is larger than the order of the polynomial.
2. Calibrating the other (P-N) positions based on the mass spectrometric data generated from the closest standard-containing position by applying the corresponding
25 polynomial transformation.

Once optimal standard positions have been identified according to the invention, mass analysis can be carried out on test analytes as follows:

- 30 1. for each standard-containing position W, a list L is computed of the sample-containing positions that are not closer to another standard position;
2. a standard position W is selected, and a mass spectrum is acquired from this position;
3. mass spectra are then acquired from the sample-containing positions in the list L corresponding to the selected standard position W, in reverse order of their distance on the target to the selected standard position W. Each spectrum is calibrated based on the

transformation computed from the selected W spectrum. An illustration is shown in Figure 4;

4. steps (2) and (3) are repeated for the next selected standard position W.

The preferred sequence of mass spectrometric data acquisition illustrated in Figure 4 is intended to normalize the amount of drift that occurs in measured mass readings. The temporal drift is reduced for the positions most remote from W that have the largest spatial offset. That is, by analyzing the farthest positions first for which the spatial offset is the largest, the temporal contribution is reduced as much as possible. Spectra from the positions with lower spatial offset (i.e., closer to W) are acquired next, presumably with higher temporal drift. Thus, the total offset for positions in L, calibrated according to standard W, is fairly constant. Furthermore, calibration may be applied periodically during analysis of a test analyte to re-acquire measurements on the standards (e.g., between two acquisitions on sample positions in a given list L) and thus further reduce temporal drift.

In a most advantageous embodiment the procedure above is applied in a layout where the standards are symmetrically deposited on the target. Preferably, the standards are such that an equivalent number of sample positions are calibrated with respect to each standard.

The above method is illustrated by taking an example of calibration as follows:

A mixture of four compounds is used as a standard and deposited on all the positions of the calibration target. The compound masses are 1046.5423, 1347.7355, 1619.8229 and 2465.2027 Daltons. It is assumed that the experimental measurements from position W_1 , a standard in the final layout, are respectively: 1046.84, 1348.09, 1620.25 and 2465.84 Daltons.

By linear regression, an affine transformation of the form $y=ax+b$ (Figure 2) is computed, where $a=0.999757$ and $b=-0.035351$. The measured masses for position L_1 , to be calibrated based on the transformation computed from the spectrum acquired from W_1 , are 1046.90, 1348.18, 1620.35 and 2466.00 Daltons (the calibration target contains the standard mixture at every position). This corresponds to the respective errors 0.36, 0.45, 0.53 and 0.80. The transformation $y=ax+b$ is then applied resulting in the L_1 calibrated masses of 1046.61, 1347.82, 1619.92 and 2465.37 Daltons. The new errors of 0.07, 0.08, 0.10 and 0.17 Daltons, or mass differences, are much smaller than the observed uncalibrated errors, and they are subsequently used in the method of the invention for the computation of the performance of position W_1 for calibrating position L_1 (e.g. to calculate the standard deviation of the statistical distribution of mass differences).

Several references to the technical literature have been made in the present disclosure. All of the referenced documents are hereby incorporated by reference.

5 The foregoing description of implementations of the invention has been presented for purposes of illustration and description. It is not exhaustive and does not limit the invention to the precise form disclosed. Modifications and variations are possible in light of the above teachings or may be acquired from practicing of the invention. The scope of the invention is defined by the claims and their equivalents.

10